

Abstract

5

10076194-034402

A method and system are disclosed for determining who is the speaking person in video data. This may be used to add in person identification in video content analysis and retrieval applications. A correlation is used to improve the person recognition rate relying on both face recognition and speaker identification. Latent Semantic Association (LSA) process may also be used to improve the association of a speaker's face with his voice. Other sources of data (e.g., text) may be integrated for a broader domain of video content understanding applications.

15